



Tutorial

José António Ribeiro Maia¹
Maria Alice Sousa¹
Rogério César Fermino¹
André Seabra¹
Simonete Silva^{1,2}
Rui Garganta da Silva¹
Alcibíades Bustamante¹
Vitor Pires Lopes³
António Prista⁴
Duarte Luís Freitas⁵

ANÁLISE E INTERPRETAÇÃO DOS NÍVEIS DE ACTIVIDADE FÍSICA DE CRIANÇAS: UM TUTORIAL BASEADO NA MODELAÇÃO HIERÁRQUICA OU MULTINÍVEL.

ANALYSIS AND INTERPRETATION OF PHYSICAL ACTIVITY LEVELS IN CHILDREN: A TUTORIAL BASED ON HIERARCHICAL OR MULTILEVEL MODELLING.

RESUMO

Este tutorial pretende apresentar, de modo didáctico, uma forma alternativa de análise de dados sobre os níveis de actividade física de crianças a partir da modelação hierárquica ou multinível. São mencionadas as diferentes etapas da modelação, os resultados são interpretados com base nos *output's* do *software* utilizado – o HLM 6.02. Em cada etapa do percurso são lançadas as hipóteses mais importantes em grau de complexidade crescente. A sua importância é referida a partir dos resultados disponíveis.

Palavras-chave: Actividade física; Crianças; Computação estatística.

ABSTRACT

This tutorial aims at a didactical presentation of an alternative approach to analyse physical activity data of children based on hierarchical or multilevel modelling. We present the basic steps of the data analysis, interpreting all relevant output from the chosen software – HLM 6.02. Each step is duly presented and explained, going from simple hypothesis to more complex ones. Their relevancy is presented in terms of the available results.

Key words: Physical activity; Children; Mathematical computing.

1 Faculdade de Desporto – Universidade do Porto.

2 Departamento de Educação Física – Universidade Regional do Cariri.

3 Instituto Politécnico de Bragança.

4 Faculdade de Ciências de Educação Física e Desporto – Universidade Pedagógica de Moçambique.

5 Universidade da Madeira – Portugal.

INTRODUÇÃO

É hoje inquestionável a importância que em todo o mundo se atribui aos níveis moderados a elevados da actividade física (ActF) e das suas implicações no bem-estar e saúde das populações. Não admira pois, que tal conhecimento e factos inquestionáveis de estudos epidemiológicos e clínicos tenham originado um forte movimento investigativo que culminou no aparecimento de uma nova área de pesquisa – a epidemiologia da ActF (EActF), formalmente estabelecida pelos textos fundamentais de Carl Caspersen^{1,2}.

Uma parte substancial da pesquisa descritiva em EActF centra a sua atenção em delineamentos de corte transversal e de natureza observacional acerca dos níveis de ActF de crianças e jovens. Este tipo de estudos amostra, geralmente, um grande número de sujeitos, normalmente na casa dos milhares. Só desta forma é possível obter estimativas precisas das estatísticas utilizadas para descrever o comportamento dos valores de ActF dos sujeitos observados, bem como salientar os efeitos pretendidos de generalização dos resultados. Formas bem ilustrativas de representar a informação são os valores médios (\pm desvio-padrão, mínimo e máximo), e os respectivos intervalos de confiança (IC). Se se dispuser de valores de corte (ainda que na maior parte dos casos arbitrários, constituídos a partir da distribuição centílica da variável em estudo), é possível calcular prevalências das diferentes categorias da variável, e que podem ser divididas, por exemplo em sedentários, moderadamente activos e vigorosos.

Uma outra corrente de investigação em EActF trata da pesquisa analítica, em que se pretende identificar os factores responsáveis pela variação observada nos valores da variável utilizada para marcar a definição abrangente de ActF. Decorre daqui a plêiade de estudos acerca dos “determinantes” (do inglês *determinants*) ou variáveis correlatas (do inglês *correlates*) que se associam às diferenças de resultados na ActF das populações infanto-juvenis (sobre esta matéria ver as revisões de Caspersen et al.², Dishman et al.³, Sallis⁴,

Pate et al.⁵, Sallis et al.⁶). Os métodos de análise são bem conhecidos, e vão desde análises de variância a regressão múltipla, ou regressão logística.

Esta forma de lidar com informação de grandes amostras, ainda que bem importante e muitas vezes altamente esclarecedora nos seus propósitos, assume que uma parte substancial das questões pode ser respondida focalizando a sua atenção na unidade mais baixa de uma larga hierarquia informacional – as crianças e os jovens. Contudo, quando observamos com cuidado a apresentação do processo amostral verificamos que na grande maioria dos casos é do tipo multi-estádio: crianças, turmas, professores, escolas, zonas geográficas, regiões distintas de um estado ou país. De facto, é bem evidente, como muito bem têm salientado autores ilustres no contexto do desempenho escolar⁷⁻¹¹ ou exclusivamente no domínio metodológico¹²⁻¹⁵, que a grande maioria da informação emergente da pesquisa epidemiológica acerca dos níveis ou padrões de ActF é de natureza hierárquica ou multinível: crianças agrupadas em turmas de características distintas, cujas aulas são leccionadas por professores de formação diversa, que estão em escolas com estruturas humanas e materiais diferenciadas, que se localizam em espaços sócio-geográficos muito distintos, etc. Estamos diante de uma grande teia informacional de natureza hierárquica que permeia um discurso de natureza ecológico mais abrangente que a simples análise da unidade mais micro desta vasta rede de inter-relações – os sujeitos.

Ora é nosso propósito apresentar uma forma alternativa de análise de dados transversais sobre a ActF de crianças com base na modelação hierárquica ou multinível (MHMN), quando a informação tem uma hierarquia própria (e praticamente toda a informação no lato universo de pesquisa da EActF tem uma estrutura multinível). Efectuaremos um percurso didáctico para auxiliar os investigadores apresentando aspectos bem relevantes desta metodologia com base no programa estatístico HLM 6.02, partindo de um problema bem concreto: o do estudo da ActF de crianças dos seis

Tabela 1. Dimensão amostral estratificada por agrupamento, ano e sexo.

Agrupamento	Nº de escolas	Número de alunos por ano e sexo								Total
		1º		2º		3º		4º		
		♂	♀	♂	♀	♂	♀	♂	♀	
Dr. Leonardo Coimbra	3	13	9	12	20	21	13	14	21	123
Vila Caíz	4	50	56	47	57	49	38	33	38	368
Amarante	17	64	69	73	76	69	63	66	72	552
Marão	15	34	34	32	37	39	33	35	37	281
São Gonçalo	18	88	109	93	111	89	103	96	102	791
Travanca	16	80	96	90	136	84	112	119	108	825
Total	73	329	373	347	437	351	362	363	378	2940
Total por ano de escolaridade		702		784		713		741		

aos 10 anos de idade, tentando analisar e interpretar não só a diferença de resultados, mas o porquê dessa diferença. Daqui que tenhamos covariáveis ou preditores ao nível dos alunos e das escolas que estes frequentam.

Informação Disponível

Os dados foram obtidos no concelho de Amarante que pertence ao distrito do Porto, que fica situado na região norte de Portugal e tem 40 freguesias distribuídas por uma área de 301,5 km². É atravessado por zonas urbanas e litorais com forte variabilidade económica. A sua rede escolar está distribuída por seis agrupamentos, com um total de 73 escolas do primeiro ciclo do ensino básico (Tabela 1).

Apesar de terem sido amostradas 92% (n=2940) das crianças dos seis aos 14 anos do universo escolar, só utilizaremos 1779 crianças de ambos os sexos dos seis aos 10 anos de idade por questões de informação omissa e baixa frequência nas idades superiores a 10 anos. Estes 1779 alunos estão distribuídos por 53 escolas, e pertencem aos quatro anos do ensino básico. A frequência de alunos por escola é bastante diversificada dadas as características demográficas e orográficas de cada freguesia.

A Figura 1 refere-se à informação descritiva mais importante dos dois níveis hierárquicos dos dados disponíveis.

A ActF foi avaliada a partir do questionário de Godin e Shephard¹⁶. É registrado o número de episódios, em média por semana, em que cada criança despende mais de 15 minutos em actividades que são classificadas como leves (3 MET's), moderadas (5 MET's) ou intensas/vigorosas (9 MET's). Os valores registrados na folha do questionário são utilizados numa equação simples para estimar a ActF realizada numa semana (ActFSemanal). O score final é obtido através da multiplicação da frequência em cada categoria pelo valor do respectivo MET e pela soma

posterior dos produtos obtidos, tal como apresentamos na equação: ActFSemanal = (3*ActF leve) + (5*ActF moderada) + (9*ActF intensa)

Todas as crianças foram entrevistadas directamente por elementos do grupo de investigação que foram previamente treinados para o efeito. As estimativas de fiabilidade obtidas a partir do procedimento de teste-reteste em 104 crianças de ambos os sexos e das quatro classes com base no coeficiente de correlação intra-classe (R) foram: meninas, R=0,90, IC_{95%}=0,72;0,97; meninos, R=0,88, IC_{95%}=0,71;0,95; todas as crianças, R=0,89; IC_{95%}=0,71;0,95. Estes valores manifestam uma elevada precisão no modo como as crianças tendem a reportar as suas percepções de ActF, tal como foi também evidente noutras pesquisas¹⁷⁻²⁰. O índice de massa corporal (IMC) foi obtido de acordo com a fórmula bem conhecida [massa corporal (kg) / estatura² (m)], e os valores de fiabilidade da massa corporal e estatura foram elevados (0,98≤R≤1,00). A determinação de aspectos relativos às condições sócio-económicas de implantação da escola, das características das infra-estruturas disponíveis, bem como do material humano disponível para leccionar aulas de educação física (EF) para as crianças estão extensamente detalhadas no livro de Sousa e Maia¹⁸.

Formalmente é incorrecto calcular médias de variáveis binárias. Contudo, não deixa de ser interessante analisar o seu valor. Por exemplo, se houvesse mais meninas que meninos (meninas foram codificadas com zero e meninos com um), a média seria inferior a zero. Se o material disponível para a prática de EF fosse muito elevado nas escolas e quase todas tivessem esses equipamentos, a média seria maior do que 0,5. Deixamos ao leitor o cuidado de interpretar o resto das estatísticas.

Análise hierárquica ou multinível

Antes de iniciarmos a nossa aventura de análise e interpretação, que começará sempre pela colocação

LEVEL-1 DESCRIPTIVE STATISTICS (relativas às crianças)					
VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
IDADE	1779	7.69	1.30	6.00	10.00
SEXO	1779	0.52	0.50	0.00	1.00
ACTF	1779	77.79	27.57	11.00	161.00
IMC	1779	16.81	2.54	10.36	39.05
LEVEL-2 DESCRIPTIVE STATISTICS (relativas ao meio escolar)*					
VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
MEIO	53	0.77	0.67	0.00	2.00
RECREIO	53	1.79	0.63	0.00	3.00
MATERIAL	53	0.02	0.14	0.00	1.00

Figura 1. Estatísticas descritivas dos dois planos da hierarquia informacional (*output* do HLM 6.02). *Infra-estruturas escolares e recursos humanos "mais relevantes" por agrupamento escolar Meio: 0=tipicamente rural; 1=rural; 2=urbano; 3=urbano; Recreio: 0=não tem; 1=espaço pequeno, mas não permite a prática de qualquer desporto; 2=espaço que permite a prática de desporto; 3=espaço que permite a prática de desporto, mais um ou dois campos de futebol de cinco; Material desportivo para a prática das aulas de Educação Física, 0=não e 1=sim).

de perguntas, atentemos na Figura 2.

É bem claro que a frequência de crianças não é a mesma em cada escola, como também é perceptível que o comportamento dos resultados da ActF não é similar ao longo da idade nestas três escolas. Na primeira parece haver um declínio na ActF, na segunda alguma estabilidade, e na terceira um incremento ao longo da idade. Esta simples evidência reclama, necessariamente, um olhar cuidadoso para o comportamento da ActF entre escolas.

Atentemos agora na Figura 3, que mostra a amplitude dos níveis de ActF de 75% das escolas aleatoriamente seleccionadas, sem as ordenar (gráfico da esquerda), ou ordenando a distribuição dos seus valores pela mediana (gráfico da direita).

Também aqui é clara a presença de um factor escola que necessariamente condiciona os valores de ActF das crianças. A ausência da sua consideração induz um forte viés na interpretação da variância dos seus resultados, bem como limita, de algum modo, a interpretação dos seus determinantes ou variáveis correlatas.

Comecemos, então, pelas diferentes etapas da MHMN, colocando em cada uma delas as questões que consideramos pertinentes e que mostrarão a

enorme potencialidade desta estrutura de pensamento e forma de analisar a informação (aspectos filosóficos e metodológicos da MHMN podem ser encontrados no texto editado por Corgeau²¹ e a elegância estatística formal e métodos computacionais são referidas em Raudenbush et al.²² e Goldstein²³). Convém desde já mencionar a ideia subjacente ao uso da MHMN: a exigência de um esforço construtivista, na medida em que se percorrem etapas de construção de modelos de complexidade crescente para interpretar o porquê da variância existente nos dois planos da hierarquia informacional²⁴⁻²⁶.

Etapa nº 1: ANOVA de efeitos aleatórios (do inglês *random effects ANOVA*)

A primeira etapa de modelação consiste em determinar a extensão da variação que existe ao nível das crianças e das escolas. Isto é, estimar o valor da variância que se situa nos dois níveis da hierarquia. Com base na ANOVA de efeitos aleatórios (também designado de modelo nulo, ou *intercept only model*) é possível obter os valores pretendidos. Este primeiro modelo é frequentemente designado de *unconditional model* por não conter qualquer covariável que explique o comportamento

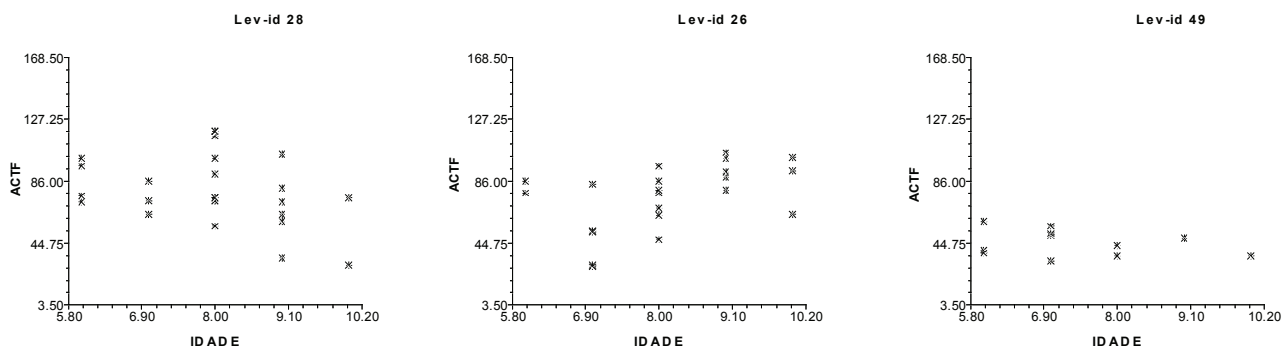


Figura 2. Distribuição dos valores de ActF em três escolas identificadas por *Lev-id* (output do HLM 6.02).

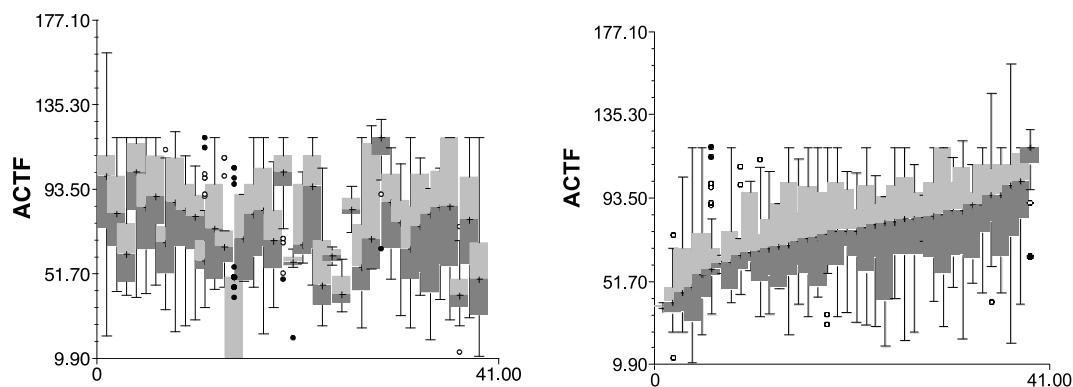


Figura 3. Amplitude dos valores da ActF de 75% das escolas (aleatoriamente escolhidas) sem e com ordenação da sua distribuição pela mediana (output do HLM 6.02).

da variância nos dois níveis. Os resultados estão no Figura 4 (ver também Figura 5), que passaremos a explicar com base no seguinte conjunto de questões:

Primeira: Qual é a média da ActF de todas as crianças amostradas, independentemente da sua idade, sexo e IMC? A grande média é de 76,16 MET's. A média populacional da ActF de todas as crianças do concelho de Amarante situa-se, com 95% de confiança, entre 72,30 e 80,02 MET's [$IC_{95\%} = 76,16 \pm 1,96(1,97)$].

Segunda: Dado que a MHMN considera variação em dois planos hierárquicos da informação, qual é a magnitude da variância entre alunos, e entre escolas? A variância inter-individual dos alunos de todas as classes e de ambos os sexos é de 544,08; a variância que marca as diferenças de ActF entre escolas é de 177,17 ($p < 0,0001$).

Terceira: Se existe uma estrutura hierárquica na informação, e o facto das crianças pertencerem a escolas distintas ser um factor diferenciador nos seus níveis de ActF, qual é, então, essa extensão? A presença de uma estrutura hierárquica nos resultados exige o cálculo do coeficiente de correlação intraclasse (ρ) que é obtido pela divisão da variância entre escolas pela variância total: $177,17 / (177,17 + 544,08) \approx 0,25$. Vinte e cinco por cento da variância total da ActF das crianças é devida a um efeito da escola, o que é um valor substancial. Esta informação é da máxima importância porque vem precisamente salientar o facto do viés interpretativo quando situamos a nossa análise exclusivamente ao nível dos alunos.

Quarta: Poderemos ter uma estimativa da fiabilidade da média da ActF das crianças e deste modo assegurar aspectos da qualidade da informação? O

valor é elevado, sendo de 0,86.

Quinta: Será que entre escolas há médias distintas de ActF? Claro que sim. Se todas tivessem a mesma média, a componente de variância (177,17) na parte aleatória do modelo (*random part* no *output* relativo ao *intercept*) não seria significativa. Mas é [$\chi^2_{(52)} = 767,37$; $p < 0,0001$]. O $IC_{95\%}$ para a ActF média das escolas é de 50,07 a 102,25 [$76,16 \pm 1,96(13,31)$], o que significa que há escolas com valores baixos de ActF e outras de valor elevado (ver Figura 5).

É mais do que evidente que a simples análise às médias a partir do modelo de efeitos fixos da ANOVA não providenciaria este manancial de informação,

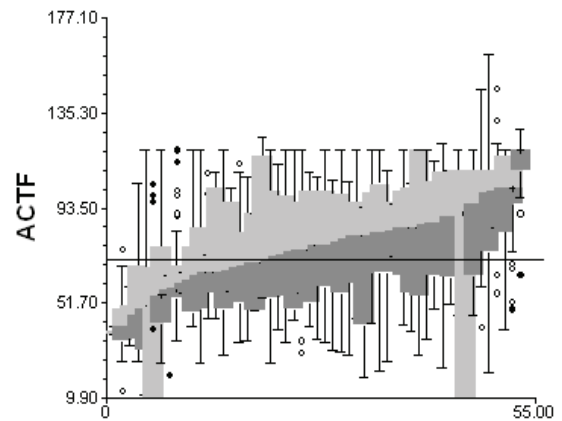


Figura 5. Amplitude da ActF de todas as escolas ordenadas pelo valor da sua mediana (*output* do HLM 6.02). A linha paralela ao eixo do "x" (i. e. das escolas) corresponde à grande média (isto é 76,17 MET's).

Random level-1 coefficient		Reliability estimate				
INTRCPT1, B0		0.859				
Final estimation of fixed effects (with robust standard errors)						
	Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For	INTRCPT1, B0 INTRCPT2, G00	76.161917	1.972472	38.612	52	0.000
Final estimation of variance components:						
Random Effect		Standard Deviation	Variant Component	df	Chi-square	P-value
INTRCPT1, level-1,	U0 R	13.31086 23.32556	177.17903 544.08187	52	767.37997	0.000
Statistics for current covariance components model						
Deviance		= 16369.527950				
Number of estimated parameters		= 3				

Figura 4. Resultados do modelo nulo (*output* do HLM 6.02)

altamente esclarecedor do quadro hierárquico e substantivo dos valores de ActF das crianças.

Centremos agora a nossa atenção num outro plano de análise:

Etapa nº 2: Modelo de efeitos fixos (do inglês *Level-1 random intercept model*)

A segunda etapa da análise implica um esforço de modelação da variância encontrada no primeiro nível da hierarquia. Isto é, ao nível das crianças, com base nas covariáveis disponíveis: idade (e aqui avançamos que com o aumento da idade também aumentam os valores de ActF); sexo (que os meninos são mais activos do que as meninas), e IMC (sendo que os que têm maiores valores são os menos activos). A escolha destes predictores faz todo o sentido, sendo que na literatura da EActF são designados de determinantes demográfico-biológicos.

Interroguemos os dados e apresentemos os principais resultados. Contudo, é importante salientar que esta etapa da MHMN obriga, provisoriamente, a que consideremos que os declives entre escolas não sejam diferentes. Isto significa que estamos a considerar a possibilidade da mesma distribuição de idades entre escolas, a mesma frequência de meninos e meninas, e a mesma distribuição de valores do IMC. Vamos então aos resultados (ver Figura 6).

Primeiro: A grande média da ActF alterou-se. E porquê? Porque o valor de referência no sexo das crianças é de zero para as meninas e um para os meninos. O valor encontrado agora; 68,90 refere-se à média de ActF das meninas.

Segundo: Os meninos são mais activos, uma vez que o valor do seu coeficiente é positivo e estatisticamente significativo ($B_2=14,20\pm 1,02$). Isto quer dizer que os meninos são, em média, mais activos do que as meninas em 14,20 MET's. As crianças mais velhas são mais activas (em média por cada ano, cerca de 2,72 MET's), e as que têm o IMC mais elevado são menos activas (-0,89 MET's por cada unidade de incremento do IMC). Estes parâmetros são estatisticamente significativos ($p < 0,01$). Estes resultados suportam as hipóteses anteriormente apresentadas.

Terceiro: É importante salientar que ainda há variância inter-individual por explicar, na ordem de 481,13. Decorre daqui que haveria que considerar outros determinantes dos alunos para diminuir aquele valor. Abrem-se portas à pesquisa para a inclusão de predictores de outras naturezas, cuja reflexão deixamos ao cuidado do leitor, convém contudo ter bem presente, como bem salientam Sallis e Owen²⁷ que para além dos factores demográfico-biológicos considerados existem outros de âmbito psicológico (motivação, gosto pelos exercícios, alcance de benefícios, desejo de exercitar-se, distúrbios do humor, percepção de saúde e aptidão senso pessoal de competência), comportamental (história de actividade anterior, qualidade dos hábitos dietéticos, processos de mudança), social e/ou cultural (influência da família e dos amigos), ambiental (acesso às facilidades, clima, custos dos programas, interrupção

da rotina) ou associado à ActF (intensidade, percepção subjectiva do esforço) com potencialidade parecem influenciar os hábitos de ActF em crianças e jovens.

Quarto: Há também variância significativa entre escolas na sua ActF, e que é estatisticamente significativa (166,68). Isto quer dizer que aspectos estruturais e do ambiente das escolas podem condicionar diferenças entre elas e ter repercussões nos valores de ActF das crianças, como tentaremos verificar mais adiante.

Quinto: Este esforço de modelação exige que se afira da sua qualidade. Passamos de um modelo não condicionado (do inglês *unconditioned model*) para um modelo condicionado (do inglês *conditioned model*), porque este último inclui predictores. Daqui que seja necessário não só informação acerca do significado estatístico dos parâmetros incluídos (e essa informação é providenciada pelo *output* e já foi anteriormente referenciada), bem como de uma estatística que nos informe acerca do valor de cada modelo. Essa estatística é designada de *deviance*. A *deviance* só faz sentido quando comparada com modelos contidos em si mesmos (do inglês *nested within*). Ora relativamente ao modelo nulo, este que agora elaboramos, o de efeitos fixos, é mais expansivo, e neste sentido pode ser comparado com o modelo nulo. Assim, a *deviance* de cada modelo é sempre apresentada no final do *output*, bem como do número de parâmetros estimados. No HLM 6.02 há uma opção de teste de hipóteses, que permite comparar dois modelos entre si desde que sejam *nested within*, isto é, hierarquicamente contidos. O resultado encontra-se numa parte designada de *model comparison test*. O resultado é estatisticamente significativo ($\chi^2_{(3)} = 215,89; p < 0,0001$). O nosso esforço de modelação foi bem sucedido.

Sexto: Qual é a variância explicada, ao nível da ActF dos alunos, pela entrada destes predictores? A resposta é bem simples. Basta para tanto contrastar as variâncias residuais entre os dois modelos, de tal modo que a variância explicada é $(544,08-481,13) / 544,08=0,116$ ($\approx 12\%$). O conjunto dos predictores idade, sexo e IMC explica 12% da variância total da ActF interindividual.

É algo difícil comparar os resultados deste estudo com os de outros autores, face à diversidade de factores e determinantes considerados. No entanto, de todas as investigações consultadas emerge a ideia, a exemplo do que aconteceu no presente trabalho, que a percentagem de variância explicada no comportamento ActF varia entre os 10 e os 30%²⁸⁻³².

Na figura 7 encontram-se as trajectórias dos sujeitos de 10 escolas. Chamamos a atenção dos leitores para as diferenças nos valores da abcissa de ambos os gráficos, que reclamam duas explicações para serem entendidos:

Primeira: Como se assumiu que o modelo era de trajectórias fixas, todas as rectas de regressão relativas à ActF entre escolas são paralelas cujo declive é necessariamente positivo, uma vez que com o aumento da idade, a ActF também aumenta.

Segunda: A idade (na abcissa) está expressa

em duas métricas distintas [o problema da centragem dos dados (do inglês *data centering*) da abcissa é um assunto que reclama uma atenção particular para a qual convidamos os leitores a um estudo atento de Hox³³, Heck e Thomas³⁴ e Raudenbush e Bryk³⁵]. Na esquerda temos os valores originais da idade dos sujeitos. Esta métrica coloca alguma dificuldade de interpretação,

dado que, conceptualmente, o valor na ordenada (47,97 MET's, valor dado num *output* não apresentado) corresponde à ActF das crianças quando a sua idade é zero, o que é uma impossibilidade. Relembramos o leitor que o *intercept* é, em língua portuguesa o valor da ordenada na origem, isto é no zero. O valor zero não tem pois qualquer significado, uma vez que à idade zero

Final estimation of fixed effects (with robust standard errors)						
Fixed Effect		Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For	NTRCPT1, B0					
	INTRCPT2, G00	68.905572	1.988029	34.660	52	0.000
For	IDADE slope, B1					
	INTRCPT2, G10	2.722258	0.952776	2.857	1775	0.005
For	SEXO slope, B2					
	INTRCPT2, G20	14.205851	1.026231	13.843	1775	0.000
For	IMC slope, B3					
	INTRCPT2, G30	-0.896968	0.350547	-2.559	1775	0.011

Final estimation of variance components:						
Random Effect		Standard Deviation	Variant Component	df	Chi-Square	P-value
INTRCPT1, level-1,	U0	12.91067	166.68527	52	847.24166	0.000
	R	21.93482	481.13631			

Statistics for current covariance components model	
Deviance	= 16153.628011
Number of estimated parameters	= 6

Model comparison test	
Chi-square statistic	= 215.89994
Number of degrees of freedom	= 3
P-value	= 0.000

Figura 6. Resultados do modelo de efeitos fixos (*output* do HLM 6.02)

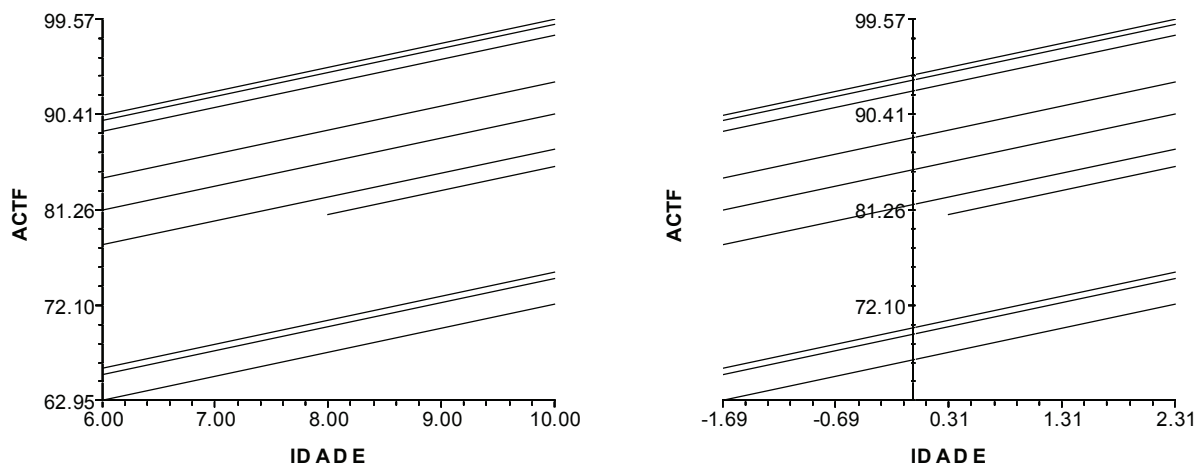


Figura 7. Rectas de regressão das primeiras 10 escolas com a idade centrada de modo distinto. No gráfico da esquerda na idade mais baixa, e no da direita na média de idades que é de 7,69 (*output* do HLM 6.02)

não pode corresponder qualquer ActF. No gráfico da esquerda o programa coloca automaticamente o eixo da ordenada no valor mais baixo de idade, isto é, seis anos. No gráfico da direita, os diferentes valores de idade de todos os sujeitos foram subtraídos à média global. O que aqui temos são, pois, dados centrados e logicamente interpretáveis, dado que o valor médio global do desempenho (o valor na ordenada) corresponde agora a uma criança média de idade igual a 7,69 anos, o que tem todo o sentido. Este valor de idade, na abcissa, corresponde ao zero, não sendo de estranhar os valores que aparecem na abcissa. Os

valores na ordenada (*intercept*) correspondem à ActF de crianças cujo valor é igual à média da sua respectiva escola. Relembramos que neste modelo, as rectas são paralelas por não se ter assumido que diferenças nas distribuições de idades, sexos e valores do IMC distinguiria a ActF entre escolas.

Etapa nº 3: Modelo de efeitos aleatórios (do inglês *Level-1 random-intercepts and random-slopes model*)

Tornando um pouco mais complexa a análise, e no sentido de explorar ainda mais o conteúdo da

Summary of the model specified (in equation format)						
Level-1 Model						
$Y = B_0 + B_1*(IDADE) + B_2*(SEXO) + B_3*(IMC) + R$						
Level-2 Model						
$B_0 = G_{00} + U_0$						
$B_1 = G_{10} + U_1$						
$B_2 = G_{20}$						
$B_3 = G_{30}$						
Final estimation of fixed effects (with robust standard errors)						
	Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For	INTRCPT1, B0					
	INTRCPT2, G00	68.781273	1.966069	34.984	52	0.000
For	IDADE slope, B1					
	INTRCPT2, G10	3.464856	0.811510	4.270	52	0.000
For	SEXO slope, B2					
	INTRCPT2, G20	14.202838	0.981801	14.466	1775	0.000
For	IMC slope, B3					
	INTRCPT2, G30	-0.897414	0.359809	-2.494	1775	0.013
Final estimation of variance components:						
	Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
	INTRCPT1, U0	12.86922	165.61682	52	837.19270	0.000
	IDADE slope, U1	3.95634	15.65265	52	147.64147	0.000
	level-1, R	21.34509	455.61301			
Statistics for current covariance components model						
	Deviance				= 16105.913154	
	Number of estimated parameters				= 8	
Model comparison test						
	Chi-square statistic				= 47.71486	
	Number of degrees of freedom				= 2	
	P-value				= 0.000	

Figura 8. Resultados do modelo com efeitos aleatórios (parte do *output* do HLM 6.02)

informação disponível, poderíamos testar as seguintes hipóteses: (1) a variação da ActF é função das diferentes distribuições de valores das idades em cada escola, (2) da distinta frequência de meninos e meninas, (3) bem como das diferentes frequências de valores de IMC entre escolas. Trata-se, agora de modelar, com maior complexidade, os aspectos aleatórios destas variáveis no contexto das suas respectivas escolas. Este esforço de modelação não permitiu suportar todas as hipóteses (*output* não mostrado). Somente se verificou plausibilidade para a primeira hipótese, e é deste modelo que temos resultados que serão interpretados (ver Figuras 8 e 9). Contudo, convém referir, uma vez mais, que estamos na presença de modelos de uma grande complexidade computacional, e que a introdução de novas covariáveis altera partes do modelo, o que mostra a sua versatilidade e sensibilidade à presença de informação adicional. De facto, a complexidade da modelação implica a adaptação do sistema de equações à novidade informativa.

Primeiro: A grande média da ActF é agora de 68,78 MET's, cuja interpretação nos é agora familiar, mas que é distinta do valor do modelo anterior, por ser condicionada em novas circunstâncias.

Segundo: O aumento da idade implica um aumento da ActF ($B_1=3,46\pm 0,81$ MET's por cada ano). Os meninos são mais activos em média 14,20 MET's ($B_2=14,20\pm 0,98$). Os que têm valores mais elevados de IMC são "penalizados" na sua ActF, em média $\approx 0,90$ MET's por cada unidade de incremento do IMC ($B_3=-0,89\pm 0,34$).

Terceiro: Ainda há variância inter-individual por explicar, e cujo valor é de 455,61; que é de menor valor do que o modelo anterior. Estão abertas as possibilidades de inclusão de outros preditores (demográfico-biológicos, psicológicos, etc) para esclarecer estas diferenças interindividuais (mais detalhes em Sallis e Owen²⁷).

Quarto: Há variância significativa no desempenho médio entre escolas, cuja magnitude é de 165,61 que é estatisticamente significativa ($p < 0,0001$).

Quinto: Ficou evidente a importância da variável idade sugerindo que a variação da sua distribuição pelas escolas implica uma atenção adequada para se entender o seu significado. Este resultado salienta aspectos diferenciadores da composição das turmas das crianças que se situam na zona montanhosa do Marão em detrimento das que se situam perto da cidade de Amarante, dadas as maiores possibilidades de fixação de agregados familiares pelas suas oportunidades económicas.

Sexto: A *deviance* deste modelo passou a ser de 16105,91; que contrastada com a do modelo anterior se revelou estatisticamente significativa ($\chi^2_{(2)}=47,71$; $p < 0,0001$), mostrando o sucesso do nosso esforço de modelação.

Sétimo: Qual é a variância explicada por este modelo? A resposta é bem simples. Basta para tanto contrastar as variâncias residuais entre os dois modelos, o nulo (Figura 4) e este, tal que a variância explicada ($544,08-455,61$) / $544,08=0,16$ (16%). Relativamente ao

modelo com efeitos fixos, ganhou-se 4%. É importante relembrar que a capacidade explicativa das variáveis consideradas neste estudo ou noutros similares é bem mais reduzida do que aquilo que é antecipado pela maioria dos investigadores (sobre esta matéria ver a revisão extensa de Sallis et al.⁶). Consideradas de modo individual, qualquer determinante explica entre cinco e 10% da variância total no melhor dos cenários de pesquisa (ver Caspersen et al.², Sallis e Owen²⁷).

Como é bem evidente na Figura 9, as rectas de regressão para as 10 escolas possuem declives distintos, salientando o efeito diferenciador dos valores de ActF em função da variabilidade na distribuição das idades das crianças.

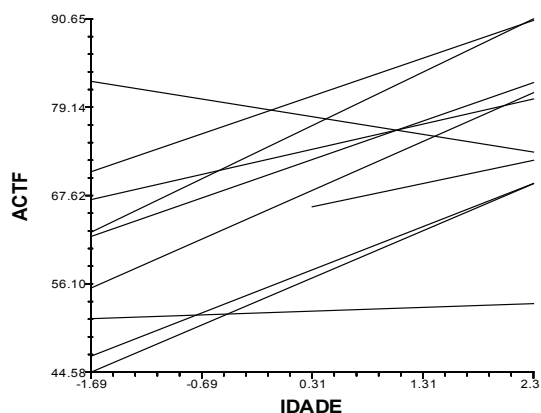


Figura 9. Rectas de regressão das primeiras 10 escolas (*output* do HLM 6.02)

Etapa nº 4: Modelo com preditores do 2º nível (do inglês *intercepts and slopes as outcomes*)

A tarefa "final" exige que nos centremos na relevância das covariáveis no plano das escolas para perceber o seu efeito ao nível distinto da ActF das crianças. Relembramos que temos três covariáveis: características do meio envolvente, do recreio e presença, ou não, de material diversificado para as aulas de EF. A inclusão destas variáveis implica a presença de interações entre os dois níveis da hierarquia informacional, que é uma riqueza analítica e interpretativa que não está disponível nos métodos mais "clássicos" de análise dos determinantes da ActF, conforme referimos na introdução.

As hipóteses que aqui lançamos são as seguintes: (1) As características do material disponível, do recreio e do meio envolvente condiciona as médias de ActF dos alunos; (2) os espaços do recreio teriam uma influência mais elevada na ActF dos mais velhos porque são os que mais "mandam" nele, e porque há aspectos de "poder infantil" de aquisição territorial face à sua idade e estatuto; (3) o recreio seria mais vantajoso para a ActF dos meninos face ao modo como tendem a ocupar os espaços disponíveis; (4) o meio (reflectindo estatuto sócio-

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B_0 + B_1*(IDADE) + B_2*(SEXO) + B_3*(IMC) + R$$

Level-2 Model

$$B_0 = G_{00} + G_{01}*(MEIO) + G_{02}*(RECREIO) + G_{03}*(MATE) + U_0$$

$$B_1 = G_{10} + G_{11}*(RECREIO) + U_1$$

$$B_2 = G_{20} + G_{21}*(RECREIO)$$

$$B_3 = G_{30} + G_{31}*(MEIO) + G_{32}*(RECREIO) + G_{33}*(MATE)$$

Final estimation of fixed effects

	Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For	INTRCPT1, B0					
	INTRCPT2, G00	56.858788	6.186747	9.190	49	0.000
	MEIO, G01	0.419515	2.935207	0.143	49	0.887
	RECREIO, G02	6.122097	3.422111	1.789	49	0.079
	MATE, G03	16.757298	12.249510	1.368	49	0.178
For	IDADE slope, B1					
	INTRCPT2, G10	8.537702	2.534544	3.369	51	0.002
	RECREIO, G11	-2.635929	1.277529	-2.063	51	0.044
For	SEXO slope, B2					
	INTRCPT2, G20	14.139471	3.959203	3.571	1767	0.001
	RECREIO, G21	0.071411	1.885405	0.038	1767	0.970
For	IMC slope, B3					
	INTRCPT2, G30	-1.194611	0.880024	-1.357	1767	0.175
	MEIO, G31	0.744417	0.345455	2.155	1767	0.031
	RECREIO, G32	-0.278355	0.446233	-0.624	1767	0.533
	MATE, G33	0.028565	0.782905	0.036	1767	0.971

Final estimation of variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1,	U0	12.09772	146.35494	49	691.21539	0.000
IDADE slope,	U1	3.67857	13.53187	51	134.29915	0.000
level-1,	R	21.31127	454.17002			

Statistics for current covariance components model

Deviance = 16090.357086

Number of estimated parameters = 16

Model comparison test

Chi-square statistic = 15.55607

Number of degrees of freedom = 8

P-value = 0.049

Figura 10. Aspecto da especificação do modelo com os três predictores (parte do *output* do HLM 6.02)

económico da implantação das escolas), os espaços do recreio e o material disponível para as aulas implicariam efeitos nos valores diferenciados do IMC e nas suas repercussões nos níveis de ActF das crianças.

Os resultados estão na Figura 10. Passaremos a interpretar somente aqueles que estão sujeitos às hipóteses anteriores, uma vez que os outros (os que se situam no primeiro plano da hierarquia) já foram explicados, se bem que neste modelo tenham valores relativamente diferentes pela entrada dos preditores deste modelo:

Primeiro: A *deviance* é menor que a do modelo anterior, implicando uma melhoria significativa marginal ($\chi^2_{(8)}=15,55$; $p=0,049$) em termos de explicação da complexidade multivariada da rede de inter-relações que se estabelecem entre os dois níveis da hierarquia.

Segundo: Os valores médios de ActF das meninas não são condicionados pelas circunstâncias do meio sócio-económico de localização da escola, dos espaços disponíveis no recreio, bem como do material escolar para as aulas de EF, uma vez que os parâmetros estimados não são estatisticamente significativos ($p > 0,05$).

Terceiro: A possibilidade da idade das crianças condicionar a exploração e ocupação dos espaços dos recreios tem um parâmetro estatisticamente significativo para esta interação ($-2,64 \pm 1,28$; $p=0,044$) não se verificando qualquer vantagem relevante dos mais velhos. Bem pelo contrário. Isto pode significar uma maior democratização da exploração e ocupação do espaço do recreio, sobretudo pelas crianças mais novas.

Quarto: Ser menino ou menina não parece condicionar a oportunidade para explorar os espaços disponíveis no recreio, uma vez que o parâmetro estimado desta interação não é significativo ($0,071 \pm 1,89$; $p=0,970$).

Quinto: Somente o meio sócio-económico de implantação da escola influencia positivamente no IMC das crianças, uma vez que o parâmetro é estatisticamente significativo ($0,74 \pm 0,35$; $p=0,031$), ampliando o seu efeito negativo nos valores de ActF das crianças.

CONSIDERAÇÕES FINAIS

Em suma, as nossas considerações situam-se em dois planos. O primeiro no que se refere ao valor estritamente conceptual desta abordagem. Daqui que esperamos que esta “viagem” guiada ao mundo da MHMN tenha evidenciado a sua elegância analítica, a riqueza enorme dos resultados que disponibiliza, mas sobretudo a exigência de um pensamento mais esclarecido e contextualizado à emergência de informação que se situa em diferentes níveis hierárquicos. O segundo ao nível empírico, em que é possível referir, genericamente, que: (1) a ActF aumenta com a idade, (2) os meninos são mais activos, (3) o IMC tem efeitos negativos na ActF das crianças, (4) a variação na frequência das idades implica efeitos distintos nos resultados de ActF em cada escola, (5) o

espaço do recreio favorece os mais novos, e que (6) o meio sócio-económico tem um efeito potenciador no IMC para reduzir os níveis de ActF das crianças dos seis aos 10 anos de idade.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Caspersen C, Powell K, Christenson G. Physical activity, exercise, and physical fitness: definitions for health-related research. *Public Health Rep* 1985;100:126-31.
2. Caspersen C, Nixon P, Durant R. Physical activity epidemiology applied to children and adolescents. *Exerc Sport Sci Rev* 1998;26:341-403.
3. Dishman R, Sallis J, Orenstein D. The determinants of physical activity and exercise. *Public Health Rep* 1985;100:158-71.
4. Sallis J. Epidemiology of physical activity and fitness in children and adolescents. *Crit Rev Food Sci Nutr* 1993;33:403-8.
5. Pate R, Long B, Heath G. Descriptive epidemiology of physical activity in adolescents. *Ped Exerc Sci* 1994;6:434-47.
6. Sallis J, Prochaska J, Taylor W. A review of correlates of physical activity of children and adolescents. *Med Sci Sports Exerc* 2000;32:963-75.
7. Boyle MH, Willms JD. Multilevel modelling of hierarchical data in developmental studies. *J Child Psychol Psychiatr* 2001;42:141-62.
8. Tinklin T. The influence of social background on application and entry to higher education in Scotland: a multilevel analysis. *Higher Education Q* 2000;54:343-85.
9. Schreiber JB, Griffin BW. Review of multilevel modelling and multilevel studies in the *Journal of Education Research* (1992-2002). *J Education Res* 2004;98:24-33.
10. Goldstein H, Rabash J, Yang M, Woodhouse G. A multilevel analysis of school examination results. *Oxford Review of Education* 1993;19:425-33.
11. Raudenbush SW, Bhumirat C. The distribution of resources for primary education and its consequences for education achievement in Thailand. *Int J Education Res* 1992;143-64.
12. Goldstein H. *Multilevel statistical models*. 3rd edition. London: Arnold; 2003.
13. Heck RH, Thomas SL. *An introduction to multilevel modeling techniques*. Mahwah: Lawrence Erlbaum Associates; 2000.
14. Raudenbush SW, Bryk A. *Hierarchical linear models. Applications and data analysis methods*. 2nd edition. Thousand Oaks: Sage Publications; 2002.
15. Zhu W. A multilevel analysis of school factors associated with health-related fitness. *Res Q Exerc Sport* 1997;68:125-35.
16. Godin G, Shephard RJ. A simple method to assess exercise behaviour in the community. *Can J Appl Sport Sci* 1985;10:141-6.
17. Maia JAR, Lopes VP, Silva RMG, Seabra A, Moraes FP, Fonseca AM, et al. Um olhar sobre crianças e jovens da região autónoma dos Açores. Implicações para a Educação Física, Desporto e Saúde. Porto: Multitema - Universidade do Porto; 2003.
18. Souza MA, Maia JAR. Crescimento somático, actividade física e aptidão física associada à saúde. Um estudo populacional nas crianças do 1º ciclo do ensino básico do concelho de Amarante. Porto: Edições da Câmara Municipal de Amarante e FADE - UP; 2005.

19. Magalhães MA. Padrão de actividade física. Estudo em crianças de ambos os sexos do 4º ano de escolaridade [Dissertação de Mestrado - Programa de Pós-Graduação em Ciências do Desporto]. Porto: Faculdade de Desporto, Universidade do Porto; 2001.
20. Sallis JF, Buono MJ, Roby JJ, Micale FG, Nelson JA. Seven-day recall and other physical activity self-reports in children and adolescents. *Med Sci Sports Exerc* 1993;25:99-108.
21. Corgeau D. Methodology and epidemiology a multilevel analysis. Approaches from different social sciences. Dordrecht: Kluwer Academic Press; 2003.
22. Raudenbush SW, Bryk A, Cheong FY, Longdon R. HLM 6.0- Hierarchical linear and nonlinear modelling. Chicago: SSI, Scientific Software International; 2004.
23. Goldstein H. Multilevel statistical models. 2nd edition. New York: Wiley & Sons, 2003.
24. Roberts JK. An introductory primer on multilevel and hierarchical linear modeling. *Learning disabilities. A Contemporary Journal* 2004;2:30-8.
25. Kreft I, DeLeuw J. Introducing multilevel modeling. Thousand Oaks; Sage Publications; 1998.
26. Snijders T, Bosker R. Multilevel analysis. An introduction to basic and advanced multilevel modeling. London: Sage Publications; 1999.
27. Sallis J, Owen N. Physical activity & Behavioral Medicine. London: Sage Publications; 1999.
28. Greendorfer SL, Lewko JH. Role of family members in sport socialization of children. *Res Q* 1978 49:146-52.
29. Wold B, Andersen N. Health promotion aspects of family and peer influences on sport participation. *J Sport Psychology* 1992;23:343-59.
30. Craig S, Goldberg J, Dietz WH. Psychosocial correlates of physical activity among fifth and eighth graders. *Prev Med* 1996;25:506-13.
31. Trost SG, Pate RR, Ward DS, Saunders R, Riner W. Correlates of objectively measured physical activity in preadolescent youth. *Am J Prev Med* 1999;17:120-6.
32. Strauss RS, Rodzilsky D, Burack G, Colin M. Psychosocial correlates of physical activity in healthy children. *Arch Pediatr Adolesc Med* 2001;155:897-902.
33. Hox J. Multilevel analysis. Techniques and applications. Mahwah: Lawrence Erlbaum Associates; 2002.
34. Heck RH, Thomas SL. An introduction to multilevel modeling techniques. Mahwah Lawrence Erlbaum Associates; 2000.
35. Raudenbush SW, Bryk A. Hierarchical linear models. Applications and data analysis methods. 2nd edition. Thousand Oaks; Sage Publications, 2002.

Endereço para correspondência

José António Ribeiro Maia
Universidade do Porto – Faculdade de Desporto
Laboratório de Cineantropometria e Gabinete de Estatística Aplicada
Rua Dr. Plácido Costa, 91 – 4200-450
Porto – Portugal.
E-mail: jmaia@fade.up.pt

Recebido em 17/04/07
Aprovado em 10/05/07